Federal Reserve Bank of New York
Staff Reports

Shared Knowledge and the Coagglomeration of Occupations

Todd M. Gabe
Jaison R. Abel

Staff Report No. 612
April 2013

Shared Knowledge and the Coagglomeration of Occupations
Todd M. Gabe and Jaison R. Abel
Federal Reserve Bank of New York Staff Reports, no. 612
April 2013
JEL classification: J24, O10, R12, R23

Abstract

This paper provides an empirical analysis of the extent to which people in different occupations locate near one another, or coagglomerate. We construct pairwise Ellison-Glaeser coagglomeration indices for U.S. occupations and use these measures to investigate the factors influencing the geographic concentration of occupations. The analysis is conducted separately at the metropolitan area and state levels of geography. Empirical results reveal that occupations with similar knowledge requirements tend to coagglomerate and that the importance of this shared knowledge is larger in metropolitan areas than in states. These findings are robust to instrumental variables estimation that relies on an instrument set characterizing the means by which people typically acquire knowledge. An extension to the main analysis finds that, when we focus on metropolitan areas, the largest effects on coagglomeration are due to shared knowledge about the subjects of engineering and technology, arts and humanities, manufacturing and production, and mathematics and science.

Key words: coagglomeration, geographic concentration, labor market pooling, knowledge spillovers, occupations

---

Gabe: University of Maine (e-mail: todd.gabe@umit.maine.edu). Abel: Federal Reserve Bank of New York (e-mail: jaison.abel@ny.frb.org). The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

I. INTRODUCTION

Geographic concentration is a fact of modern economic life. Cities and their surrounding metropolitan areas are the ultimate manifestations of the strong agglomerative forces that provide benefits associated with the close proximity of people and businesses. Indeed, metropolitan areas in the United States now account for more than 80 percent of the population and 90 percent of the country's output (Abel and Gabe, 2011). For many goods and services, a large share of total output is made in relatively few places. Some of the best known and widely-cited examples of industry agglomeration in the United States are high technology manufacturing firms located in and around Silicon Valley (Saxenian, 1994), the production of textiles in the southeast (Krugman, 1991; Ellison and Glaeser, 1997), and casino gaming in Las Vegas (Kolko, 2010).

Explanations about why industries agglomerate have evolved from Marshall's (1920) classic ideas about the benefits of a pooled labor force, the availability of intermediate inputs, and knowledge spillovers; to the presence of natural advantages for particular industry-location combinations (Ellison and Glaeser, 1999); to Duranton and Puga's (2004) microfoundations of agglomeration arising from sharing, matching, and learning externalities. Empirical studies of agglomeration use a variety of explanatory variables as proxies for these different factors, and examine their impacts on geographic concentration using measures such as the locational Gini coefficient (Krugman, 1991; Audretsch and Feldman, 1996; Jensen and Kletzer, 2006) and the Ellison-Glaeser industry concentration index (Rosenthal and Strange, 2001; Lu and Tao, 2009). More recently, the emphasis has changed from analyzing patterns of agglomeration to coagglomeration in order to shed light on the factors influencing the geographic concentration of industries (Duranton and Overman, 2005, 2008; Ellison, Glaeser, and Kerr, 2010; Kolko; 2010; Jofre-Monseny, Marín-López, Viladecans-Marsal, 2011).

As a natural extension to this research, this paper provides what we believe is the first empirical analysis of the extent to which people in different occupations locate near one another, or what we term *occupational coagglomeration*.[1] To do so, we present and examine measures of the coagglomeration of U.S. occupations at the state and metropolitan area levels. Importantly, our analysis covers the full spectrum of the U.S. economy. By contrast, most empirical studies of industry agglomeration focus on the manufacturing sector (Barrios et al., 2004; Devereux, Griffith, and Simpson, 2004; Duranton and Overman, 2005; Ellison, Glaeser, and Kerr, 2010), which accounts for less than ten percent of U.S. employment. Kolko (2010) broadened the analysis of industry concentration to include the service sector, which revealed some interesting differences in the patterns of industry location. Thus, to fully capture the geographic concentration of economic activity, it is important to move beyond specific sectors of the economy.

We then turn our attention to identifying factors underlying the patterns of occupational coagglomeration that are identified, with a focus on the importance of the similarity of knowledge required to perform a job. Reorienting the focus from industries to occupations changes the way that we think about the determinants of agglomeration. Industries are defined along the basis of "what firms make" (e.g., good or service produced), while occupations are organized by "what people do" (e.g., skills and

---

[1] To illustrate just how new studies of coagglomeration are to the literature, Helsley and Strange (2012) note that the term itself does not even appear in the most recent *Handbook of Regional and Urban Economics* (Henderson and Thisse, eds., 2004). Our occupational-based approach draws on other studies that have focused on the skills- and knowledge-based content of work (Gabe, 2009; Bacolod, Blum, and Strange, 2009a, 2009b; Scott, 2009; Abel and Gabe, 2011).

knowledge requirements) in their jobs (Feser, 2003). This means that, for example, whereas input-output relationships—characterizing the amount of one good needed to produce another—might influence the settlement patterns of some firms, this determinant of industry agglomeration is less relevant to the study of occupations. Less constrained by production relationships that dictate how things are made, people are apt to locate around others involved in the same types of work activities (e.g., computer programming), thinking less about whether their peers are employed by companies making similar or different types of goods and services. Thus, we expect occupations with similar knowledge profiles to exhibit strong patterns of coagglomeration.

Such a pattern of geographic location facilitates movement among jobs and provides a constant market for skill. The idea here, as conceived by Marshall (1920), is that workers seek out places that provide the best chance of employment and mitigate against the employment swings of companies caused by random demand shocks (Overman and Puga, 2010). In the case of occupations, workers benefit from locating in places with an abundance of jobs that require the knowledge they possess as well as jobs with similar knowledge requirements. For example, someone with high knowledge about economics, math, accounting, and computing would likely locate in a place with ample job opportunities for economists, as well as employment opportunities in related occupations such as budget analysts, cost estimators, or financial analysts. A constant market for skill in occupations with similar knowledge requirements—facilitating labor mobility among jobs—is a benefit to workers that would likely contribute to a high coagglomeration of such occupations (Fallick, Fleishman, and Rebitzer, 2006; Freedman, 2008).

A second reason why occupations with similar knowledge requirements are apt to coagglomerate is because such a locational pattern facilitates information sharing among workers. The idea of a knowledge spillover is that workers benefit from being able to learn “the mysteries of the trade” from a high concentration of similar workers. In the context of occupations, this means that knowledge intensive jobs are likely to exhibit higher levels of agglomeration and that occupations with similar knowledge profiles are apt to show strong patterns of coagglomeration. Just as an environmental economist can benefit by learning from other environmental economists, he or she can be more productive by learning from environmental engineers or wildlife ecologists. These jobs, although different in some respects, are similar in that they require high levels of knowledge regarding ecology, sciences, and the environment.

More generally, these benefits can be thought to arise from the knowledge that is shared between people in different occupations. In the case of labor pooling, sharing in the same knowledge as others in the local labor market allows workers to move between jobs more easily. In the case of knowledge spillovers, being around people with similar knowledge facilitates the sharing of information and ideas. Indeed, because these types of spillovers are often transmitted through the movement of workers between companies (e.g., inter-firm mobility of engineers within regions as examined by Almeida and Kogut (1999)), Combes and Duranton (2006) argue that labor market pooling and knowledge spillovers cannot be viewed as distinct motives for agglomeration. As such, we characterize these benefits as arising through *shared knowledge*.

To assess whether people in occupations with shared knowledge locate around one another, we estimate regression models of the determinants of occupational coagglomeration at the metropolitan area and state levels of geography. This approach of analyzing multiple spatial units draws on studies of industry agglomeration by Rosenthal and Strange (2001) and Kolko (2010), among others, and allows us to shed light on the underlying forces contributing to agglomeration. Further, to address potential endogeneity concerns, we develop an instrumental variables approach that relies on similarities in the way people typically acquire knowledge—education and experience—to instrument for the knowledge that is shared among occupations. Not only is our proposed instrument set a strong predictor of this shared knowledge, but it is plausible that any effect on the coagglomeration of occupations operates only through the knowledge that is shared by people in different occupations.

Empirical results reveal that occupations with similar knowledge requirements tend to coagglomerate and, importantly, that the effect of shared knowledge on occupational coagglomeration is about twice as large in metropolitan areas as in states. Taken together, these findings are consistent with other research showing that the benefits of labor market pooling and knowledge spillovers tend to be highly localized (Jaffe, Trajtenberg, and Henderson, 1993; Rosenthal and Strange, 2001, 2008; Fu, 2007; Jofre-Monseny, Marín-López, Viladecans-Marsal, 2011). An extension to the main analysis, focusing on different types of knowledge, finds that shared knowledge about subjects such as engineering and technology or arts and humanities have a stronger influence on coagglomeration than shared knowledge about health services or transportation. Overall, this research offers a new way to view the geographic concentration of economic activity and, as a result, provides additional insight into the determinants of agglomeration.

II. COAGGLOMERATION OF U.S. OCCUPATIONS

Following Ellison and Glaeser (1997) and Ellison, Glaeser, and Kerr (2010), we compute pairwise coagglomeration indices to assess the extent of occupational coagglomeration at the metropolitan area and state levels of geography.[2] The index for the coagglomeration of occupations *k* and *l* is:

$$OccCoaggl_{k,l} = \Omega \,/\, (1 - \sum_{i=1}^{n} t_n^2) \qquad (1)$$

where, $\Omega = \sum_{i=1}^{n} (s_{i,k} - t_i)(s_{i,l} - t_i)$

i = U.S. metropolitan areas (n=283) or states (n=51)

$s_{k(l)}$ = metro area's (state's) share of employment in occupation k (l)

t = metro area's (state's) share of total employment.

With 468 occupations available, this index is calculated for 109,278 distinct occupational pairs using IPUMS data from the 2010 American Community Survey of the U.S. Census Bureau (Ruggles et al. 2011). Positive index values suggest that the occupations are both agglomerated in the same places, index values of close to zero indicate no tendency to coagglomerate, and negative index values suggest that the occupations are both agglomerated but in different places.[3]

---

[2] Data limitations prevent us from examining occupational coagglomeration at the zip code or county levels. However, given the nature of our analysis, we believe metropolitan areas are an appropriate geographic unit of observation because they best represent the labor market areas in which workers interact and move among jobs. Our reliance on occupational pairs as the unit of observation necessarily excludes any within-occupation variation that may exist.

[3] Another indicator of coagglomeration, proposed by Duranton and Overman (2005, 2008), incorporates information on the distances between plants operating in industry pairs. We are unable to calculate the Duranton and Overman metric of coagglomeration for occupations because, although the region (i.e., state and metropolitan area) of employment is known, the exact location is not.

Largely by construction, the mean value of this coagglomeration index is approximately zero. While many occupational pairs are not coagglomerated, the index exhibits considerable dispersion, indicating that some occupations do tend to agglomerate in the same places. At the metropolitan area level, the index values range from -0.022 to 0.070, with a standard deviation of 0.002. At the state level of analysis, the index values range from -0.036 to 0.136, with a standard deviation of 0.003. When compared to parallel measures of industry coagglomeration, the extent of occupational coagglomeration appears to be on par with that observed in the service industries, but less than what is observed for manufacturing industries (Ellison, Glaeser, and Kerr, 2010; Kolko, 2010).

The highest occupational coagglomeration pairs in U.S. metropolitan areas and states are shown in Tables 1 and 2, respectively. Looking at Table 1, we see that occupations exhibiting the strongest patterns of coagglomeration in U.S. metropolitan areas are those involved in casino gaming (e.g., gaming services workers, gaming cage workers, and gaming managers), television and motion pictures (e.g., actors, producers and directors, camera operators and editors, and agents), and matters related to the economy and analysis of businesses (e.g., economists, operations research analysts, budget analysts, information security analysts). The strong patterns of coagglomeration in occupations related to gaming and the television-motion picture production are not surprising. As noted by Kolko (2010), the gaming (Las Vegas) and television-motion picture (Los Angeles) sectors are two of the most geographically concentrated industries, and the occupations exhibiting high coagglomeration are closely related to these sectors.

The high levels of coagglomeration among economists, operations research analysts, budget analysts, and information security analysts is somewhat less expected, given that these occupations are spread across many types of industries, but these patterns are consistent with the notion that shared knowledge among occupations may result in a high coagglomeration of occupations. The occupations that tend to co-locate with economists, with the exception of astronomers and physicists, have similar job requirements that need high levels of knowledge pertaining to subjects such as economics and business, mathematics, budgeting and finance, as well as strong computing and analytical skills. Although budget analysts, operations research analysts, and economists work across a variety of industries—at least, more so than gaming and television-motion picture workers—the required skill set can be easily transferred across these jobs.

An examination of the occupational pairs shown in Table 2 suggests that the jobs exhibiting strong patterns of coagglomeration are very different when focusing on states rather than metropolitan areas. The occupations with the strongest patterns of coagglomeration in U.S. states are involved in textiles manufacturing (e.g., textile winding, textile knitting, and textile bleaching) and extraction (e.g., petroleum and mining engineers, drill operators, pumping station operators, and geological and petroleum technicians). These occupations are included among the most coagglomerated at the state level, but not when examining the coagglomeration of occupations at the metropolitan area level—due to their location across entire states (e.g., textiles in North Carolina and Georgia). The absence of gaming and "economics-related" occupations in Table 2, given their prominence in Table 1, is testament to the fact that these occupations tend to coagglomerate in certain metropolitan areas (e.g., gaming-related occupations in

Las Vegas and Atlantic City; economists and related occupations in Washington DC and New York), but the high levels of co-location rarely expand to elsewhere in the state.

Although some of the television and motion picture-related occupations prominent in Table 1 also make an appearance in Table 2, they include some "unusual pairings" in the list of highest coagglomeration pairs examined at the state level. For example, whereas actors are coagglomerated with producers and directors, camera operators, and agents in metropolitan areas, they are connected to agricultural workers and health practitioners—along with some of the same television-motion picture-related occupations as in Table 1—in states. Although the results of high patterns of coagglomeration in metropolitan areas reflect the fact that actors, directors and camera operators work together to make television shows and motion pictures, the finding of a high state-level coagglomeration among actors and agricultural workers (and, to a lesser extent, health practitioners) is merely an artifact that some of the same states (e.g., California and Florida) with robust entertainment industries also tend to have large agricultural sectors.

This descriptive analysis provides some intuition about the factors that might influence the coagglomeration of occupations in the United States. For example, people working in occupations that tend to produce the same output (e.g., casino gaming, television and motion pictures, textiles) appear to locate in the same places. And, in many cases, these places tend to be at distinct points along the urban spectrum (e.g., big cities). However, it is also clear that many of the most concentrated occupations share similar knowledge requirements (e.g., economists, operations research analysts, budget analysts). Moreover, because occupational coagglomeration differs in metropolitan areas and states, the importance of these factors may depend on geography. With this in mind, we now turn to a more formal analysis of the determinants of occupational coagglomeration.

III. SHARED KNOWLEDGE AND OCCUPATIONAL COAGGLOMERATION

To investigate whether shared knowledge influences the coagglomeration of occupations in the United States, we estimate a regression model that examines the relationship between the coagglomeration indexes described above and a new measure of the knowledge that is shared across occupations. Specifically, using distinct occupational pairs as observations, we estimate the following model:

$$OccCoaggl_{k,l} = \alpha + \beta SharedKnowledge_{k,l} + \phi \boldsymbol{X}_{k,l} + \mu_{k,l} + \varepsilon_{k,l} \qquad (2)$$

where k,l denotes an occupational pair, $\boldsymbol{X}$ is a vector of controls, μ is fixed effect indicating whether the two occupations are part of the same major occupational category, and ε is an error term.

This is the same estimation approach set forth by Rosenthal and Strange (2001) to analyze the determinants of industry agglomeration and used by Ellison, Glaeser, and Kerr (2010) and Kolko (2010) to examine industry coagglomeration. We perform our analysis using coagglomeration indexes calculated at the metropolitan area and state levels of geography. In order to facilitate comparisons in the coefficient estimates obtained from the regression analysis, all of the variables except for the same major occupation indicator are standardized to have a mean value of zero and a standard deviation of 1.0. Table 3 presents definitions and data sources for the variables used in our regression analysis, while a more detailed description of these variables is provided below.

A. Explanatory Variables

The explanatory variable of key interest, *Shared Knowledge*, is used to measure the similarity of knowledge required between occupations. This variable is calculated using information from the U.S. Department of Labor's Occupational Information Network (O*NET), which is collected from interviews of incumbent workers and the input of occupational analysts.[4] We focus our attention on the knowledge required to perform a job in the 33 subject areas, which are combined into 10 subject groups, as shown in Table 4. Data on these knowledge requirements are collected using a two-part question that asks: (1) the importance of a knowledge area to a job (on a scale of 1 to 5), and (2) the level of knowledge required (on a scale of 1 to 7) in cases where a knowledge area is determined to be at least somewhat important (a score of 2 or higher on the first question). For each of the 468 occupations included in the analysis, we calculated a knowledge index (*KI*) that is the product of a job's knowledge importance multiplied by its knowledge level.[5]

Shared Knowledge, shown in equation 3, is a measure of the similarity of knowledge requirements for occupations *k* and *l*:

$$Shared\ Knowledge_{k,l} = -\sum_{z=1}^{33} (KI_{k,z} - KI_{l,z})^2 \tag{3}$$

where the subscript *z* indicates the subject area and *KI* is the knowledge index. Because higher values of this variable indicate a greater similarity in the knowledge profiles of occupations, we expect to find a positive relationship between it and the coagglomeration index.

4 See Peterson et al. (2001) for a detailed discussion of O*NET.

5 This is similar to the approach used by Feser (2003) and Abel and Gabe (2011).

Along with the similarity in the types of knowledge required for a job, we expect other factors to influence occupational coagglomeration patterns. Given the often strong connection between certain types of occupations and industries, it is likely that jobs contributing to the same industry (e.g., actors, directors, camera operators, costume designers) will exhibit strong patterns of coagglomeration. Ellison, Glaeser, and Kerr (2010) and Kolko (2010) find that the similarity in the types of occupations employed—used as a proxy for the importance of labor market pooling—contributes to higher levels of industry coagglomeration. Similarly, we expect that occupations involved in producing the same types of goods and services will exhibit stronger tendencies to coagglomerate.

Similar Output, shown in equation 4, is a measure of the extent to which the distribution of employment across major industrial categories is similar among occupations.

$$\textit{Similar Output}_{k,l} = -\sum_{j=1}^{19} (IS_{k,j} - IS_{l,j})^2 \tag{4}$$

where the subscript j indicates the major industrial category and IS is the share of U.S. occupational employment, obtained from the 2010 American Community Survey, in the industry. High values of this variable suggest that workers in the two occupations make similar goods and services, while low values indicate that workers in the occupations contribute to different sectors. As such, we expect this variable to have a positive effect on occupational coagglomeration, suggesting that occupations contributing to similar industries are more likely to concentrate in the same places.

Another control variable used in the analysis, *Similar City Size*, seeks to account for similarities among jobs in the city-size distribution of occupations. Indeed, Kolko

(2010) argues that service industries tend to urbanize more than manufacturing industries in large part because they rely less on natural resources. Thus, just as the availability of natural resources dictates where certain types of industries are located, the presence of a large population can influence the location patterns of some occupations. Explaining the role of shared natural advantages in the coagglomeration of industries, Ellison, Glaeser, and Kerr (2010) note that even in the absence of other benefits of agglomeration, certain types of industries that rely on, say, a coastal location will coagglomerate. A similar argument can be made for the role of the city-size distribution of employment on occupational coagglomeration patterns. As an example of what we have in mind, for a variety of reasons, a high percentage of professional athletes, subway train conductors, and chief executives of Fortune 500 companies work in large cities. These occupations generally require a different set of skills and they largely contribute to different industries. The only common thread among these occupations is that their jobs are commonly found in big cities, just as—in the example of industry coagglomeration—certain sectors are connected through nothing more than a common reliance on the same natural resource.

Similar City Size, shown in equation 5, represents the extent to which the two occupations exhibit similar employment distributions across the urban population hierarchy.

$$\textit{Similar City Size}_{k,l} = -\sum_{s=1}^{7} (POP_{k,s} - POP_{l,s})^2 \qquad (5)$$

where the subscript s indicates the population size category and *POP* is the share of U.S. occupational employment, obtained from the 2010 American Community Survey, in metropolitan areas included in the size category. For each of the 468 occupations, we calculated the share of employment in seven metropolitan area population size categories, including an option that indicates a non-metropolitan location.[6] Occupations such as Economists, Actors, Producers and Directors, Agents and Business Managers of Artists, and Financial Analysts tend to have higher shares of employment in places with 5 million or more people, while Tire Builders, Farmers and Ranchers, Logging Workers, Mining Machine Operators, and Explosives Workers are almost nonexistent in the largest metropolitan areas. We expect this variable—which takes on high values when the occupations have similar population size distributions—to have a positive effect on coagglomeration. Such a relationship would be consistent with the idea that some occupational pairs may be concentrated in the same places simply because they both tend to locate in metropolitan areas with similar population size distributions.

B. Baseline Empirical Results

Table 5 presents OLS regression results on the determinants of occupational coagglomeration in U.S. metropolitan areas and states. Overall, the empirical models perform reasonably well. The R-squared values of almost 0.10 are similar to the goodness-of-fit measures reported in comparable studies of industry agglomeration and coagglomeration (Rosenthal and Strange, 2001; Ellison, Glaeser, and Kerr; 2010; Gabe and Abel, 2012). Further, all of the explanatory variables are statistically significant at the 1-percent level.

Consistent with theories of labor market pooling and knowledge spillovers, the regression results show that, other things being equal, shared knowledge between

[6] The seven population size categories are: over 5 million people, between 2.5 million and 5 million people, between 1 million and 2.5 million people, between 500,000 and 1 million people, between 250,000 and 500,000 people, less than 250,000 people, and non-metropolitan.

occupations is positively related to occupational coagglomeration. In the analysis of coagglomeration at the metropolitan area level, a one-standard deviation increase in the Shared Knowledge variable is associated with a 0.095-standard deviation increase in the value of the coagglomeration index. The estimated coefficient corresponding to the Shared Knowledge variable in the analysis of coagglomeration at the state level is about one-half of this size (0.051), indicating that the effect of knowledge similarity on coagglomeration is larger at a more intimate level of geographic analysis. In other words, having similar knowledge requirements matters more to coagglomeration at the metropolitan area level than it does at the state level.

These findings are consistent with other research demonstrating that the benefits of shared knowledge tend to be highly localized (Jaffe, Trajtenberg, and Henderson, 1993; Rosenthal and Strange, 2001, 2008; Fu, 2007; Jofre-Monseny, Marín-López, Viladecans-Marsal, 2011, among others). This is because moving between jobs with similar knowledge requirements is easier and less costly within a labor market than between them. Likewise, the physical proximity that exists at smaller spatial scales helps to facilitate the flow of knowledge by increasing the amount of interaction and face-to-face contact that people experience (Storper and Venables, 2004; Abel, Dey, and Gabe, 2012).

Turning to other regression results, the estimated coefficients corresponding to the Similar Output variable, which suggest that occupations contributing to the same types of industries have stronger tendencies to coagglomerate, are relatively similar at the two levels of geographical analysis. The impact of the Similar Output variable is less than one-half of the magnitude of the impact of the Shared Knowledge variable in the analysis of coagglomeration at the metropolitan area level, which suggests that the similarity of the occupation's knowledge profile is relatively more important than the similarity of the types of goods and services produced at this level of geography. However, when looking at state-level coagglomeration patterns, the impacts of the Shared Knowledge and Similar Output variables are about the same.

The Similar City Size variable also shows roughly the same relationship with occupational coagglomeration at the metropolitan area and state levels of analysis. In both cases, the estimated coefficients corresponding to the metropolitan area size variable are indicative of larger associations than those ascribed to similarities in the knowledge requirements of occupations. The regression results also suggest that belonging to the same major occupational category contributes to the value of the coagglomeration index. Other things being equal, the coagglomeration index for metropolitan areas and states increases by 0.262 and 0.367 standard deviations, respectively, in cases where the two occupations belong to the same major SOC category.

C. Instrumental Variables Estimation

The possibility that locational patterns of occupations could influence the types and similarities of knowledge that are required in a job might raise concerns about our baseline OLS estimation. Instrumental variables estimation provides a strategy to address this type of identification problem. Indeed, focusing on industries, Ellison, Glaeser, and Kerr (2010) employ such an approach to mitigate the concern that "industrial relationships may be the result of co-location instead of the cause of co-location." However, implementing instrumental variables estimation requires that we identify variables that are correlated with the knowledge that is shared between occupations (i.e., relevant) but not directly related to their degree of coagglomeration (i.e., exogenous).

To construct our instrument set, we focus on the means by which people acquire knowledge to explain similarities in the knowledge profiles among occupations. We use information on an occupation's required education and experience—specifically, the similarity of these ways of obtaining knowledge among occupation pairs—as instruments for the Shared Knowledge variable. For the education variable, the O*NET survey includes a set of 12 response categories such as "less than a high school diploma," "bachelor's degree," and "first professional degree," while the experience variable includes 11 response categories such as "up to and including 1 month" and "over 8 years, up to and including 10 years." Using information on the percentage of O*NET survey respondents who selected each category, we calculated the *Similar Education* and *Similar Experience* variables as the sum of the squared differences between the occupational pairs. The logic of the instrument set is that similar knowledge profiles likely arise from the same types of education or experiences; however, to the extent there is a relationship between these variables and the geographic concentration of occupations, it occurs only through the similarity in knowledge required to perform a job.

First-stage regression results presented in Table 6 indicate that the instrument set is a strong predictor of the knowledge shared among occupations. To assess the strength of these instruments, we used the Stock and Yogo (2005) weak instrument test that compares the first-stage F-statistic to a critical value that depends on the number of endogenous variables, the size of the instrument set, and the tolerance for the "size distortion" of a test (α=0.05) of the null hypothesis that the instruments are weak. We can reject the null hypothesis of weak instruments based on the Stock and Yogo (2005) test using a 10-percent maximal size threshold.

With the relevance criterion satisfied, we now consider the exogeneity of the instrument set. Our key identifying assumption here is that any relationship between the coagglomeration of occupations and the means by which people acquire knowledge occurs though the Shared Knowledge variable. That is, firms and workers locate around a specific skill set and, in general, do not consider how the skills were acquired (Marshall, 1920; Krugman, 1991). Thus, we believe it is plausible that our instrument set is exogenous. Consistent with this idea, over-identification test results, with *p*-values of 0.504 and 0.691, indicate that the instrument set is uncorrelated with the error terms. As our instrument set satisfies the relevance and exogeneity conditions, we conclude that the instruments are valid.

Second-stage regression results presented in Table 6 suggest that the Shared Knowledge variable has a positive and significant effect on the occupational coagglomeration index when potential endogeneity is taken into account. The estimated coefficient corresponding to the Shared Knowledge variable is 0.167 at the metropolitan area level and 0.102 at the state level, which is considerably higher than the estimated coefficients from the OLS regressions. The difference between the IV and OLS estimates suggests that there may be some measurement error in the Shared Knowledge variable. Overall, though, findings from the instrumental variables estimation diminish concerns that the baseline results are being driven by endogeneity between patterns of occupational coagglomeration and the knowledge that is shared among occupations.

D. Empirical Results by Type of Knowledge

The baseline OLS results and IV estimation suggest that the similarity of knowledge required across a wide range of topics is a key determinant of occupational coagglomeration, especially when studying metropolitan areas. As an extension to this analysis, we examine how shared knowledge about specific subjects influences the locational patterns of occupations. To do so, we use the knowledge area groupings shown in Table 4, and recalculate 10 versions of the Shared Knowledge variable—each one uses information on the individual knowledge area(s) that are included in the group. For example, shared knowledge about engineering and technology is calculated as the sum of the squared differences in the knowledge indices for computers and electronics, engineering and technology, design, building construction, and mechanical.

Table 7 shows OLS regression results on the relationship between shared knowledge and occupational coagglomeration patterns for U.S. metropolitan areas and states, by subject. To illustrate, Figure 1 presents the relative magnitudes of the estimated coefficients that correspond to the standardized values of the shared knowledge variables. Two estimated coefficients are shown in the figure for each of the knowledge subjects: one for metropolitan areas (left) and one for states (right). Focusing on the estimates for metropolitan areas, we see that the largest effect on coagglomeration due to shared knowledge is from engineering and technology, followed by arts and humanities, manufacturing and production, and mathematics and science. Shared knowledge about health services has no effect on occupational coagglomeration for metropolitan areas, and the estimated coefficients corresponding to shared knowledge about transportation, law and public safety, and education and training are relatively small.

Comparing the estimated coefficients from the analysis of metropolitan areas and states illustrates some interesting patterns related to the forces influencing occupational coagglomeration. It is apparent from Figure 1 that the effects of shared knowledge on coagglomeration are larger in metropolitan areas than states for the subjects of engineering and technology, mathematics and science, arts and humanities, and education and training. This means that relative close proximity enhances the benefits—either through knowledge spillovers or a pooled labor force—from being around others with similar knowledge about these topics. On the other hand, the effects of shared knowledge on coagglomeration are more similar between metropolitan areas and states for the subjects of manufacturing and production, business and management, and communications.

Figure 1 also reveals that the order of importance of the knowledge subjects on occupational coagglomeration differs between metropolitan areas and states. For states, engineering and technology has the largest estimated coefficient—similar to what we found for metropolitan areas—but it is followed in magnitude by business and management, communications, and manufacturing and production. Shared knowledge about mathematics and science, and arts and humanities is relatively unimportant for state-level occupational coagglomeration, despite their relatively large effects when examining metropolitan areas.

The empirical results shown in Table 7 and Figure 1 are in line with other studies about externalities arising from particular types of economic activity. In particular, our findings showing the importance of shared knowledge about engineering and technology to occupational coagglomeration are consistent with the study by Almeida and Kogut

(1999), who found that the movement of engineers among firms in a region, especially Silicon Valley, contributes to knowledge spillovers. Gabe and Abel (2011) uncovered high levels of geographic concentration in the knowledge-based occupational clusters of social scientists, engineers, scientists and artists, which is similar to our results showing positive effects on (metropolitan area) occupational coagglomeration due to shared knowledge about engineering and technology, arts and humanities, and mathematics and science. Finally, our results pertaining to education and training, and health services—indicating relatively small (or no) effects from shared knowledge on coagglomeration—are similar to those reported in a study by Abel, Dey, and Gabe (2012) showing that the industry sector of education and health is not characterized by substantial exchanges of information or sharing of ideas.

IV. CONCLUSIONS

Understanding the factors that influence the geographic concentration of economic activity is at the core of urban economics and regional science. Alfred Marshall's (1920, p. 225) classic ideas about labor market pooling and knowledge spillovers—which suggest that workers seek out places "where there are many employers who need such skill as theirs" and benefit from being able to learn "the mysteries of the trade" from being around similar workers—emphasize the strong connection between worker skills and geographic concentration. Focusing on similarities in the knowledge requirements across a wide variety of topics, this paper presents new evidence on the importance of shared knowledge to the geographic concentration of economic activity.

To do so, we focus our attention on the extent to which people in different occupations locate near one another, or coagglomerate. As such, we construct new measures of occupational coagglomeration at the state and metropolitan area levels of geography using information on the full spectrum of the U.S. economy. In contrast, previous studies on the geographic concentration of industries have generally focused on sectors within a major industrial sector, such as manufacturing. Similar to past research focusing on industries, we find that many occupations do tend to agglomerate in the same places.

We then use the measures of coagglomeration along with information on the similarities of occupations to examine factors that may contribute to the observed geographic patterns of where jobs are located. This requires a different way of thinking about the forces of agglomeration. Although some occupations are closely linked to specific industries (e.g., textile machine workers and the textiles industry), many occupations (e.g., executives, clerical workers, computer technicians) cut across all sectors of the economy. Moreover, whereas factors such as input-output relationships might influence the settlement patterns of some firms, they are less relevant to the study of occupations and where people locate. Instead, people are apt to locate around others involved in the same types of work activities, thinking less about whether their peers are employed by companies making similar or different types of goods and services.

Consistent with theories of labor market pooling and knowledge spillovers, empirical results reveal that occupations with similar knowledge requirements tend to coagglomerate. Importantly, we also demonstrate that the effect of this shared knowledge on occupational coagglomeration is about twice as large in metropolitan areas as in states. These findings are robust to instrumental variables estimation that relies on an instrument set characterizing the means by which people typically acquire knowledge. An extension to the main analysis shows that, when focusing on metropolitan areas, the largest effects on coagglomeration are due to shared knowledge about the subjects of engineering and technology, arts and humanities, manufacturing and production, and mathematics and science. Overall, these findings provide new evidence on the importance of shared knowledge as a factor influencing the geographic concentration of economic activity.

REFERENCES

Abel, Jaison R. and Todd M. Gabe. 2011. "Human Capital and Economic Activity in Urban America," *Regional Studies,* 45, 1079-1090.

Abel, Jaison R., Ishita Dey, and Todd M. Gabe. 2012. "Productivity and the Density of Human Capital," *Journal of Regional Science*, 52, 562-586.

Almeida, Paul and Bruce Kogut. 1999. "Localization of Knowledge and the Mobility of Engineers in Regional Networks," *Management Science*, 45, 905-917.

Audretsch, David and Maryann Feldman. 1996. "R&D Spillovers and the Geography of Innovation and Production," *American Economic Review*, 86, 630-640.

Bacolod, Marigee, Bernardo S. Blum, and William C. Strange. 2009a. "Skills in the City," *Journal of Urban Economics,* 65, 136-153.

Bacolod, Marigee, Bernardo S. Blum, and William C. Strange. 2009b. "Urban Interactions: Soft Skills vs. Specialization," *Journal of Economic Geography*, 9, 227-262.

Barrios, Salvador, Luisito Bertinelli, Eric Strobl, and Antonio-Carlos Teixeira. 2005. "The Dynamics of Agglomeration: Evidence from Ireland and Portugal," *Journal of Urban Economics*, 57, 170-188.

Combes, Pierre-Philippe and Gilles Duranton. 2006. "Labour Pooling, Labour Poaching, and Spatial Clustering," *Regional Science and Urban Economics*, 36, 1-28.

Devereux, Michael, Rachel Griffith, and Helen Simpson. 2004. "The Geographic Distribution of Production Activity in the UK," *Regional Science and Urban Economics*, 34, 533-564.

Duranton, Gilles and Henry Overman. 2005. "Testing for Localization Using Micro-Geographic Data," *Review of Economic Studies*, 72, 1077-1106.

Duranton, Gilles and Henry Overman. 2008. "Exploring the Detailed Location Patterns of U.K. Manufacturing Industries using Microgeographic Data," *Journal of Regional Science*, 48, 213-243.

Duranton, Gilles and Diego Puga. 2004. "Micro-Foundations of Urban Agglomeration Economies," Chapter 48 in Henderson, J.V., Thisse, J.F. (eds.), *Handbook of Regional and Urban Economics*, Vol. 4, Amsterdam: Elsevier/North-Holland, pp. 2063-2117.

Ellison, Glenn and Edward L. Glaeser. 1997. "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," *Journal of Political Economy*, 105, 889-927.

Ellison, Glenn and Edward L. Glaeser. 1999. "The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration," *American Economic Review*, 89, 311-316.

Ellison, Glenn, Edward L. Glaeser, and William R. Kerr. 2010. "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns," *American Economic Review*, 100, 1195-1213.

Fallick, Bruce, Charles Fleischman, and James Rebitzer. 2006. "Job-Hopping in Silicon Valley: Some Evidence Concerning the Microfoundations of a High-Technology Cluster," *Review of Economics and Statistics*, 88, 472-481.

Feser, Edward. 2003. "What Regions Do Rather than Make: A Proposed Set of Knowledge-Based Occupation Clusters," *Urban Studies*, 40, 1937-1958.

Freedman, Matthew. 2008. "Job Hopping, Earnings Dynamics, and Industrial Agglomeration in the Software Publishing Industry," *Journal of Urban Economics*, 64, 590-600.

Fu, Shihe. 2007. "Smart Café Cities: Testing Human Capital Externalities in the Boston Metropolitan Area," *Journal of Urban Economics*, 61, 86-111.

Gabe, Todd M. 2009. "Knowledge and Earnings," *Journal of Regional Science*, 49, 439-457.

Gabe, Todd M. and Jaison R. Abel. 2011. "Agglomeration of Knowledge," *Urban Studies*, 48, 1353-1371.

Gabe, Todd M. and Jaison R. Abel. 2012. "Specialized Knowledge and the Geographic Concentration of Occupations," *Journal of Economic Geography*, 12, 667-691.

Helsley, Robert W. and William C. Strange. 2012. "Coagglomeration and the Scale and Composition of Clusters," unpublished manuscript.

Jaffe, Adam, Manuel Trajtenberg, and Rebecca Henderson. 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*, 108, 577-598.

Jensen, J. Bradford and Lori Kletzer. 2006. "Tradable Services: Understanding the Scope and Impact of Services Offshoring," in Brainard, L., Collins, S. (eds.), *Brookings Trade Forum 2005: Offshoring White-Collar Work*, Washington DC: Brookings Institution Press, pp. 75-133.

Jofre-Monseny, Jordi, Raquel Marín-López, and Elisabet Viladecans-Marsal. 2011. "The Mechanisms of Agglomeration: Evidence from the Effect of Inter-Industry Relations on the Location of New Firms," *Journal of Urban Economics*, 70, 61-74.

Kolko, Jed. 2010. "Urbanization, Agglomeration, and Co-Agglomeration of Service Industries," Chapter 5 in Glaeser, E.L. (ed.), *Agglomeration Economics*, Chicago: University of Chicago Press for the NBER, 151-180.

Krugman, Paul. 1991. *Geography and Trade*, Cambridge, MA: The MIT Press.

Lu, Jiangyong and Zhigang Tao. 2009. "Trends and Determinants of China's Industrial Agglomeration," *Journal of Urban Economics*, 65, 167-180.

Marshall, Alfred. 1920. *Principles of Economics,* London: Macmillan.

Overman, Henry and Diego Puga. 2010. "Labour Pooling as a Source of Agglomeration: An Empirical Investigation," Chapter 4 in Glaeser, E.L. (ed.), *Agglomeration Economics,* Chicago: University of Chicago Press for the NBER, 133-150.

Peterson, Norman, Michael Mumford, Walter Borman, Richard Jeanneret, Edwin Fleishman, Kerry Levin, Michael Campion, Melinda Mayfield, Frederick Morgeson, Kenneth Pearlman, Marilyn Gowing, Anita Lancaster, Marilyn Silver, and Donna Dye. 2001. "Understanding Work Using the Occupational Information Network (O*NET): Implications for Practice and Research," *Personnel Psychology*, 54, 451-492.

Rosenthal, Stuart S. and William C. Strange. 2001. "The Determinants of Agglomeration," *Journal of Urban Economics*, 50, 191-229.

Rosenthal, Stuart S. and William C. Strange. 2008. "The Attenuation of Human Capital Spillovers," *Journal of Urban Economics*, 64, 373-389.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew D. Schroeder, and Matthew Sobek. 2010. *Integrated Public Use Microdata Series: Version 5.0* [Machine-readable database]. Minneapolis: University of Minnesota.

Saxenian, AnnaLee. 1994. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*, Cambridge, MA: Harvard University Press.

Scott, Allen. 2009. "Human Capital Resources and Requirements across the Metropolitan Hierarchy of the USA," *Journal of Economic Geography*, 9, 207-226.

Stock, James H. and Motohiro Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression," in Andrews, D.W.K., Stock, J.H. (eds.), *Identification and Inference for Economic Models: A Festrschrift in Honor of Thomas Rothenberg*, Cambridge: Cambridge University Press, 80-108.

Storper, Michael and Anthony Venables. 2004. "Buzz: Face-to-Face Contact and the Urban Economy," *Journal of Economic Geography,* 4, 351-370.

Table 1. 20 Highest Coagglomeration Pairs, U.S. Metropolitan Areas

Occupation 1	Occupation 2	Co-Agglomeration
Economists	Operations Research Analysts	0.070
Gaming Services Workers	Gaming Cage Workers	0.064
Producers and Directors	Actors	0.061
Gaming Managers	Gaming Services Workers	0.061
Gaming Managers	Gaming Cage Workers	0.060
Sewing Machine Operators	Actors	0.059
Geological and Petroleum Technicians, and Nuclear Technicians	Petroleum, mining and geological engineers	0.057
Television, Video, and Motion Picture Camera Operators and Editors	Actors	0.054
Miscellaneous Social Scientists	Economists	0.053
Textile Knitting and Weaving Machine Setters, Operators, and Tenders	Textile Winding, Twisting, and Drawing Out Machine Setters, Operators, and Tenders	0.046
Miscellaneous extraction workers including roof bolters and helpers	Petroleum, mining and geological engineers	0.045
Budget Analysts	Economists	0.045
Information Security Analysts	Economists	0.045
Derrick, rotary drill, and service unit operators, and roustabouts, oil, gas, and mining	Petroleum, mining and geological engineers	0.043
Petroleum, mining and geological engineers	Chemical Engineers	0.042
Agents and Business Managers of Artists, Performers, and Athletes	Actors	0.040
Astronomers and Physicists	Economists	0.039
Miscellaneous textile, apparel, and furnishings workers	Actors	0.038
Environmental Scientists and Geoscientists	Petroleum, mining and geological engineers	0.038
Sewing Machine Operators	Producers and Directors	0.034

Notes: Coagglomeration index is from Ellison, Glaeser, and Kerr (2010). Data source is the 2010 American Community Survey of the U.S. Census Bureau, accessed using IPUMS-USA (Ruggles et al., 2011).

Table 2. 20 Highest Coagglomeration Pairs, U.S. States

Occupation 1	Occupation 2	Co-Agglomeration
Textile Winding, Twisting, and Drawing Out Machine Setters, Operators, and Tenders	Textile Knitting and Weaving Machine Setters, Operators, and Tenders	0.136
Derrick, rotary drill, and service unit operators, and roustabouts, oil, gas, and mining	Petroleum, mining and geological engineers	0.121
Miscellaneous extraction workers	Petroleum, mining and geological engineers	0.106
Petroleum, mining and geological engineers	Pumping Station Operators	0.095
Textile Winding, Twisting, and Drawing Out Machine Setters, Operators, and Tenders	Textile bleaching and dyeing, and cutting machine setters, operators, and tenders	0.089
Derrick, rotary drill, and service unit operators, and roustabouts, oil, gas, and mining	Miscellaneous extraction workers	0.082
Geological and Petroleum Technicians, and Nuclear Technicians	Petroleum, mining and geological engineers	0.078
Actors	Graders and Sorters, Agricultural Products	0.076
Derrick, rotary drill, and service unit operators, and roustabouts, oil, gas, and mining	Pumping Station Operators	0.074
Miscellaneous extraction workers	Pumping Station Operators	0.067
Model Makers and Patternmakers, Metal and Plastic	Electronic Equipment Installers and Repairers, Motor Vehicles	0.066
Television, Video, and Motion Picture Camera Operators and Editors	Actors	0.062
Textile bleaching and dyeing, and cutting machine setters, operators, and tenders	Textile Knitting and Weaving Machine Setters, Operators, and Tenders	0.062
Producers and Directors	Actors	0.058
Television, Video, and Motion Picture Camera Operators and Editors	Graders and Sorters, Agricultural Products	0.054
Derrick, rotary drill, and service unit operators, and roustabouts, oil, gas, and mining	Geological and Petroleum Technicians, and Nuclear Technicians	0.054
Petroleum, mining and geological engineers	Miscellaneous Plant and System Operators	0.053
Actors	Health Diagnosing and Treating Practitioners	0.052
Textile Winding, Twisting, and Drawing Out Machine Setters, Operators, and Tenders	Miscellaneous textile, apparel, and furnishings workers	0.051
Miscellaneous extraction workers	Geological and Petroleum Technicians, and Nuclear Technicians	0.050

Notes: Coagglomeration index is from Ellison, Glaeser, and Kerr (2010). Data source is the 2010 American Community Survey of the U.S. Census Bureau, accessed using IPUMS-USA (Ruggles et al., 2011).

Table 3. Variable Definitions and Data Sources

Variable Name	Definition	Data Source
Occupational Coagglomeration	Ellison, Glaeser, and Kerr (2010) coagglomeration index, calculated at metro area and state level of geography	2010 ACS
Shared Knowledge	Similarity among occupations in required knowledge, examined across 33 subject areas	O*NET
Similar Output	Similarity among occupations in the share of workers by major industrial category	2010 ACS
Similar City Size	Similarity among occupations in the share of workers in 7 metropolitan population size categories (e.g., less than 250,000, over 5 million)	2010 ACS
Major Occupation	=1 if occupations are in the same major SOC category; 0 otherwise	2010 ACS
Similar Education	Similarity among occupations in the share of workers in 12 educational categories (e.g., high school diploma, associate's degree, post-master's certificate)	O*NET
Similar Experience	Similarity among occupations in the share of workers in 11 experience categories (e.g., none; over 1month, up to and including 3 months; over 2 years, up to and including 4 years)	O*NET

Table 4. O*NET Knowledge Areas and Subject Groups

Business and Management	*Mathematics and Science*	*Arts and Humanities*
Administration and Management	Mathematics	English Language
Clerical	Physics	Foreign Language
Economics and Accounting	Chemistry	Fine Arts
Sales and Marketing	Biology	History and Archeology
Customer and Personal Service	Psychology	Philosophy and Theology
Personnel and Human Resources	Sociology and Anthropology	
	Geography	*Law and Public Safety*
Manufacturing and Production		Public Safety and Security
Production and Processing	*Health Services*	Law and Government
Food Production	Medicine and Dentistry	
	Therapy and Counseling	*Communications*
Engineering and Technology		Telecommunications
Computers and Electronics	*Education and Training*	Communications and Media
Engineering and Technology	Education and Training	
Design		*Transportation*
Building and Construction		Transportation
Mechanical		

Note. The 33 O*NET knowledge areas (shown in plain text) are combined into 10 subject groups (shown in italics).

Table 5. OLS Results: Determinants of Occupational Coagglomeration

Variable	Estimated Coefficients Metro-Level	 State-Level
Constant	-0.016*	-0.023*
	(0.003)	(0.003)
Shared Knowledge	0.095*	0.051*
	(0.003)	(0.003)
Similar Output	0.046*	0.059*
	(0.003)	(0.003)
Similar City Size	0.249*	0.232*
	(0.004)	(0.005)
Major Occupation	0.262*	0.367*
	(0.018)	(0.019)
R-squared	0.091	0.082
Number of Observations	109,278	109,278

Notes: Robust standard errors are shown in parentheses. * denotes statistical significance at the 1-percent level.

Table 6. IV Results: Determinants of Occupational Coagglomeration

	Estimated Coefficients			
	Metro-Level		State-Level	
Variable	First-Stage	Second-Stage	First-Stage	Second-Stage
Constant	-0.043*	-0.012*	-0.043*	-0.020*
	(0.003)	(0.003)	(0.003)	(0.003)
Shared Knowledge	--	0.167*	--	0.102*
		(0.007)		(0.007)
Similar Output	0.053*	0.042*	0.053*	0.057*
	(0.003)	(0.003)	(0.003)	(0.003)
Similar City Size	0.002	0.245*	0.002	0.223*
	(0.002)	(0.004)	(0.002)	(0.005)
Major Occupation	0.680*	0.193*	0.680*	0.318*
	(0.008)	(0.019)	(0.008)	(0.020)
Similar Education	0.433*	--	0.433*	--
	(0.003)		(0.003)	
Similar Experience	-0.001	--	-0.001	--
	(0.003)		(0.003)	
Partial R-squared	0.193	--	0.193	--
F-statistic for Stock-Yogo Strong Instrument Test	10,972.95**	--	10,972.95**	--
P-value for Over-Identification Test	--	0.504	--	0.691
Number of Observations	109,278	109,278	109,278	109,278

Notes: Robust standard errors are shown in parentheses. * denotes statistical significance at the 1-percent level. IV estimates obtained using limited information maximum likelihood (LIML) estimator. ** denotes we can reject the null hypothesis of weak instruments based on the Stock-Yogo test (α=0.05) using the 10% maximal LIML size threshold (e.g., critical value of 8.68 in a model with 2 instruments).

Table 7. OLS Results: Shared Knowledge and Occupational Coagglomeration, by Subject

Variable	Estimated Coefficients: Metro-Level	State-Level
Constant	-0.016*	-0.022*
	(0.003)	(0.003)
Shared Knowledge about:		
Business and Management	0.028*	0.029*
	(0.003)	(0.003)
Manufacturing and Production	0.032*	0.026*
	(0.002)	(0.002)
Engineering and Technology	0.062*	0.035*
	(0.003)	(0.003)
Mathematics and Sciences	0.030*	0.002
	(0.004)	(0.004)
Health Services	-0.002	0.007
	(0.003)	(0.003)
Education and Training	0.013*	0.003
	(0.002)	(0.002)
Arts and Humanities	0.036*	0.014*
	(0.005)	(0.005)
Law and Public Safety	0.009*	0.006*
	(0.002)	(0.002)
Communications	0.027*	0.028*
	(0.005)	(0.004)
Transportation	0.008*	0.014*
	(0.003)	(0.003)
Similar Output	0.053*	0.062*
	(0.003)	(0.003)

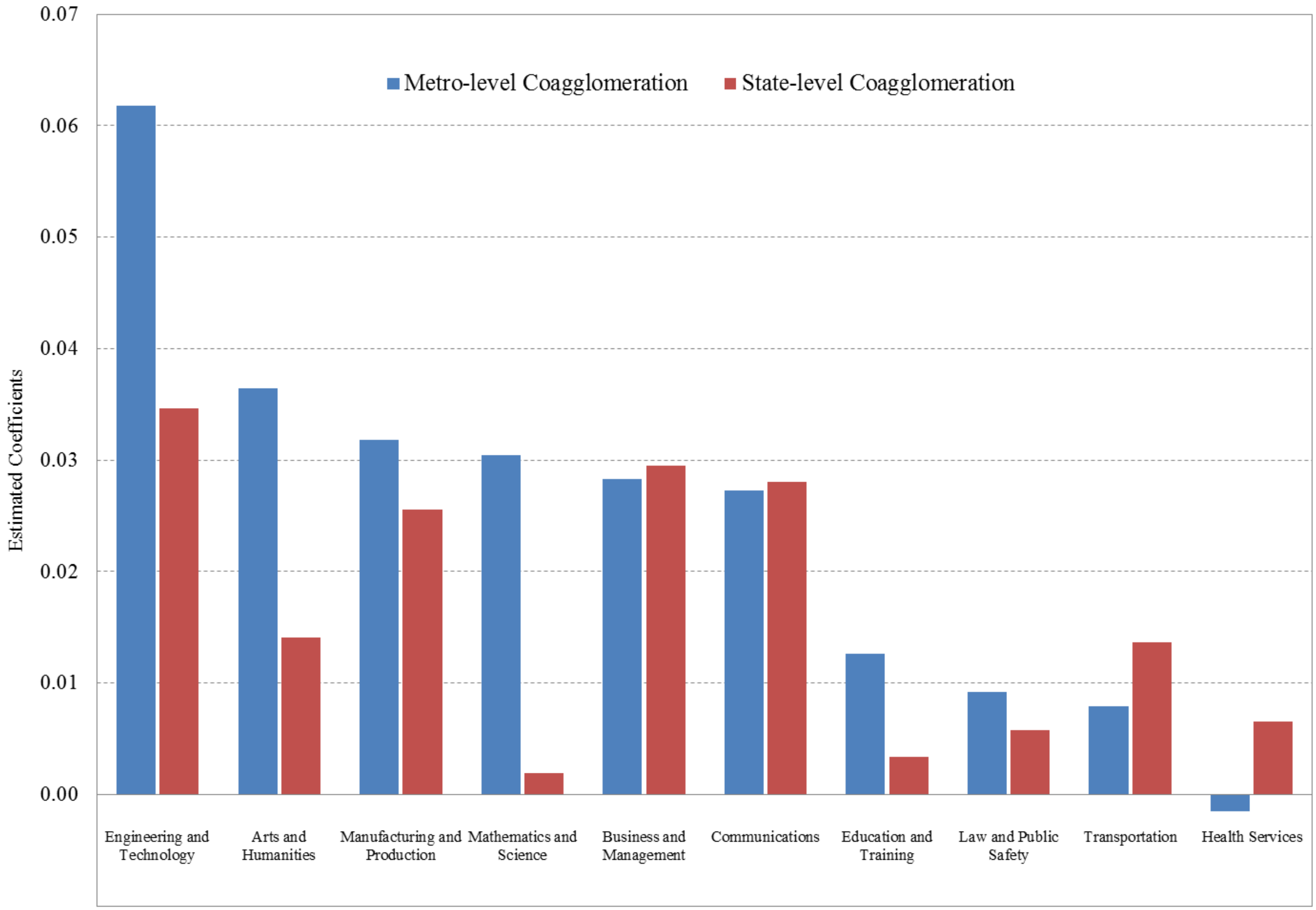
Table is continued on the following page.

Table 7. OLS Results: Shared Knowledge and Occupational Coagglomeration, by Subject

	Estimated Coefficients	
Variable	Metro-Level	State-Level
Similar City Size	0.239*	0.225*
	(0.004)	(0.005)
Major Occupation	0.249*	0.353*
	(0.018)	(0.019)
R-squared	0.093	0.084
Number of Observations	109,278	109,278

Notes: Robust standard errors are shown in parentheses. * denotes statistical significance at the 1-percent level.

Figure 1. Shared Knowledge and Occupational Coagglomeration, by Subject



Note: Estimated coefficients for each subject and geography are reported in Table 7.